

**APPLICATION**  
**FOR**  
**UNITED STATES LETTERS PATENT**

APPLICANT NAME FREDRICK K. P. ~~K~~LASSEN

TITLE

System and Method for Determining Network Throughput Speed and  
Streaming Utilization

DOCKET NO. END9 2000 0102 US1

**INTERNATIONAL BUSINESS MACHINES CORPORATION**

**CERTIFICATE OF MAILING UNDER 37 CFR 1.10**

I hereby certify that, on the date shown below, this correspondence is being  
deposited with the United States Postal Service in an envelope addressed to  
the Assistant Commissioner for Patents, Washington, D.C., 20231 as  
"Express Mail Post Office to Addressee" on 12/21/00

Mailing Label No. EL598673206US

Name of person mailing paper: Denise M. Jurik

Denise M Jurik 12/21/00  
Signature Date

**SYSTEM AND METHOD FOR DETERMINING NETWORK THROUGHPUT SPEED  
AND STREAMING UTILIZATION**

**Background of the Invention**

Cross References to Related Applications

5 U.S. patent application Serial No. 08/\_\_\_\_\_, assignee  
docket number END9 2000 0184, entitled "SYSTEM AND METHOD  
FOR DETERMINING NETWORK DISCRETE UTILIZATION" filed  
concurrently herewith is assigned to the same assignee  
hereof and contains subject matter related to the subject  
10 matter of the present application. The above identified  
patent application is incorporated herein by reference.

This application is an improvement upon copending U.S.  
patent application Serial No. 09/267,843, filed 12 March  
1999, by F. K. P. Klassen and R. M. Silverman, entitled  
15 "SYSTEM AND METHOD FOR ANALYZING AND TUNING A COMMUNICATIONS  
NETWORK", and S/N 09/452,403 filed 1 Dec 1999 by K. Chan, F.  
K. P. Klassen, and R. M. Silverman for "System and Method  
for Monitoring Performance, Analyzing Capacity and  
Utilization, and Planning Capacity for Networks and

Intelligent, Network Connected Processes".

### Technical Field of the Invention

5 This invention relates to testing and analyzing a communications network. More particularly, the invention relates to determining throughput speed and streaming utilization.

### Background Art

10 Users and managers of communications networks are interested in the capacity, performance, reliability, availability, topology, service level attainment, and modeling of their networks. Currently, many separate, incompatible, complicated, and often unsatisfactory tools are required to perform the tasks required by network  
15 managers.

One prior art technique provides for communication system analysis in which sequences of pings of different sizes are sent, transmitted in isolation from each other, in

order to determine characteristics of the network beyond availability of resources and "usual" response time. In accordance with this technique, the best and mean times for the sequences of pings of different lengths are measured, and then the following are geometrically derived via use of the reciprocal of the packet size and bandwidth slope:

- the network path's bandwidth,
- the network path's propagation delay, and
- the network path's queuing time.

Further network testing involves sending bursts of pings to determine the loss rate of packets, from which the "internal packet size of the network" can then be determined. By this it is meant that the minimum device buffer size in the transmission path of the network can be discovered. The sending of bursts of pings of decreasing size is also described, the intent of which is to determine if smaller packets are prioritized over larger packets. In this way, the following is determined:

- the network path's minimum buffer size (also referred

to as "internal packet size"); and

- the network's short frame prioritization capability.

It is recognized in the art that there will generally be a  
5 plurality of network hops between the test station and the  
remote station, each with its own bandwidth. In these  
circumstances, taking the reciprocal of the slope of the  
line as the bandwidth is equivalent to saying that:

$$(1 / \text{Bandwidth}) = (1 / b_1 + 1 / b_2 + \dots 1 / b_n)$$

10 where  $b_1, b_2, \dots b_n$  are the bandwidths of all the hops in  
the path. This will produce a bandwidth close to accurate  
only if there is one hop in the path that is significantly  
slower than all the others. Where this is not the case, the  
estimate may be inaccurate because no account is taken of  
15 the fact that for a multi-hop path, the intermediate nodes  
will delay onward transmission of a packet until the whole  
packet has been received.

US Patent Applications S/N 09/267,843 filed 12 Aug 1999  
by F. K. P. Klassen and R. M. Silverman for System and  
20 Method for Analyzing and Tuning a Communications Network

(Klassen and Silverman), and S/N 09/452,403 filed 1 Dec 1999 by K. Chan, F. K. P. Klassen, and R. M. Silverman for System and Method for Monitoring Performance, Analyzing Capacity and Utilization, and Planning Capacity for Networks and  
5 Intelligent, Network Connected Processes (Chan, Klassen and Silverman) improved upon the then existing art by:

- distinguishing between and calculating the values for the streaming and discrete speeds of the network; and
- calculating the utilization of that connection.

10 The distinction between the discrete and streaming speeds of a network allows for two accurate measures of network speed across multiple hops, as opposed to the previously known one, potentially inaccurate measure.

The discrete speed reflects the network's maximum  
15 bandwidth for handling "conversational" type transaction traffic, such as query-response. The streaming speed reflects the network's maximum speed for "batch" type traffic, such as file transfers and print. The network's streaming speed can also be referred to as "network  
20 throughput bandwidth."

Thus, probative test packets are transmitted, received and time stamped, and then mathematically analyzed. The result is a system and method in which full network capacity analysis can be performed across networks of any size, including the Internet, and to do so with no need for network documentation or knowledge of network topology or knowledge of capacity of network components.

Prior art methods for determining the streaming speed (the maximum end-to-end speed for file and other batch transfers) include the following.

First, transmission of a large stream of echo (or one way) traffic. The bits transmitted are divided by total transmission time to determine the network's bidirectional (or one way) streaming speed.

Second, transmission of a burst of echo (or one way) traffic. The number of bits transmitted is divided by total transmission time to determine the network's bidirectional (or one way) streaming speed.

Third, transmission of a burst of long packets and a burst of an equal number of short packets. Then, (total

long bits in burst minus total short bits in burst) is divided by (long burst transmission time minus short burst transmission time).

By "transmission of a burst of packets," is meant that  
5 a plurality of packets, for example ten packets, is transmitted in as rapid succession as possible. For calculations of network speed, a number of such burst tests or large stream tests are performed. The calculations for (maximum) network speed are then based upon the minimum  
10 transmission time experienced, and the calculation for average network speed, useful for utilization measures, is based upon the average transmission time experienced.

Each of these methods for determining streaming speed suffers from certain deficiencies. Unless performed at a  
15 time when there is absolutely no traffic anywhere on the connection, transmission of a stream of data interferes with existing user traffic and is interfered with by existing user traffic. In all other cases, such testing is disruptive to the network and the results are rendered  
20 inaccurate. On occasion, such testing can be useful in a small, private network. In most large networks, and across the Internet, such testing is both inaccurate and



undesirable.

Unlike these streaming techniques, prior art methods for determining network speed derived from burst or discrete speed are not disruptive to the network. However, like the  
5 above streaming techniques, they are highly subject to inaccuracy, albeit for a different reason. While large streams of traffic will fill the network, bursts of traffic often will not fill the network, especially over connections that are long or fast. So current testing techniques  
10 relying on measuring total transmission and receipt time for a burst, or relying on comparison of transmission and receipt times for bursts of different size packets, will measure both productive and dead time on the network, and will, therefore, arrive at false calculations of network  
15 streaming speed. Heretofore, all probative test measure methods for determination of end to end network speed for multihop connections suffered from inaccuracy and, in the case of streaming, disruptiveness as well, especially in large network environments and in Internet connections.

20 An alternative method for determining streaming speed lies somewhere between bursting and streaming, and is referred to as a complex burst test. In the complex burst

test method, a burst of long messages and a burst of short messages are transmitted, and the difference in bits sent is divided by the difference in transmission time in order to determine the streaming speed of the network. This method  
5 is accurate only if the bursts fill the network, and is otherwise prone to error. This problem is exacerbated by longer network lengths and faster networks.

It is an object of the invention to provide an improved system and method for network measurement and planning.

10 It is a further object of the invention to provide an improved method for calculating the streaming speed of a network.

It is a further object of the invention to provide the new concept that in a multihop network, there are two  
15 different measures of utilization, namely, streaming utilization and discrete utilization.

It is a further object of the invention to provide an improved method for determining end to end network streaming utilization.

It is a further object of the invention to provide an improved method for tuning a network.

It is a further object of the invention to provide an improved method for estimating the current performance of a network and its future performance under changed conditions ("what-if" scenarios).

### Summary of the Invention

In accordance with the invention, a system and method is provided for evaluating a communications network. A plurality of network evaluation signals, or probative test packets, are selectively sent and received through the network. Responsive to these evaluation signals, selective network evaluation parameters are determined and stored. Queuing theory analysis, responsive to these parameters, determines the response time and throughput characteristics, including streaming capacity, utilization and performance, of the network.

Other features and advantages of this invention will

become apparent from the following detailed description of the presently preferred embodiment of the invention, taken in conjunction with the accompanying drawings.

### **Brief Description of the Drawings**

5           Figure 1 illustrates a networked system in accordance with the preferred embodiment of the system of invention.

          Figure 2 illustrates a communications network and network evaluation system in accordance with a preferred embodiment of the system of the invention.

10           Figure 3 illustrates a representative logic flow for user input.

          Figure 4 illustrates the logic flow of the application and network throughput routine of the preferred embodiment of the invention.

15           Figure 5 geometrically illustrates ping and transaction results and bandwidth, utilization, and message size

derivations in accordance with the preferred embodiment of the invention.

### Best Mode for Carrying Out the Invention

Referring to Figure 1, a communications network cloud  
5 20 is depicted with a client station 21, a server station 25  
(or they can be peer stations) and a network management  
station 23 connected to the cloud at its boundary. The  
program code embodying a preferred embodiment of the  
invention can reside in any one or in any combination in  
10 each of the three stations 21-25. As used herein, a  
networked system refers to a network 20, intelligent end  
stations 21, 25, and the network 20 together with its  
intelligent end stations 21, 25 considered as a unitary  
whole.

15 Referring to Figure 2, in accordance with the preferred  
embodiment of the invention, apparent network speed analysis  
(ANSA) application 30 executes on workstation 22 to measure,  
monitor, estimate, capacity plan, and tune communications  
network 20 with respect to target station 24. A user input  
20 device 26 and data store 28 are provided at work station 22,

and ANSA 30 includes a throughput routine 34, a service level and capacity planning routine 36, and optimized ping and other routines 38.

Throughput analysis routine 34 provides for defining,  
5 calculating, and using the following network concepts: Hop  
Count Factor, Duplex Factor, Throughput Factor, and  
Multi-Server Factor. Routine 34 relates to the receiving  
station, which as shown here may also be the sending  
station. As will be more fully described hereafter, in an  
10 exemplary embodiment, routines 38 first determine the MTU  
size of the network between send station 22 and target  
station 24, and then sends bursts of that packet size to  
target station 24. Target station 24 may echo or not. If  
it echos, throughput routine 34 would typically be provided  
15 at work station 22. However, routine 34 may also be  
provided at some station other than the station originating  
the packet bursts. If target station 24 does not echo, then  
it must time stamp the bursts and either do a throughput  
routine itself or provide them to some other station to  
20 perform the throughput analysis.

Routine 34 analyzes the time stamps to determine  
throughput speed of the network, which the line speed,

which is the inverse of slope of line (e,i) (Figure 5) and calculated as the total bits in a burst packet divided by the time from receipt of the first bit to the last bit.

Routine 34 also includes a throughput utilization routine,  
5 deriving the throughput utilization as  $(1 - (\text{current available speed} / \text{maximum throughput speed}))$ .

Service level and capacity planning routine 36,  
responsive to the skilled use of routines 32 and 34,  
provides comprehensive "what-if" network planning  
10 facilities; calculation of the increase (or change) in  
network traffic before network response time service level  
is compromised; calculation of the additional file load  
capacity of the network, which is the additional file load  
before response time is compromised (for both prioritized  
15 and non-prioritized network cases); and determination of  
tuning recommendations for recommended window size for file  
transfer to fill remaining capacity (both respective and  
irrespective of maintaining response time service level, and  
for both prioritized and non-prioritized cases). In  
20 addition, routine 36 performs comprehensive what if end  
processor planning for the end systems with respect to  
processor-bound, I/O-bound, and general transactions.

Optimized ping and other methodologies routines 38 provides for optimized pinging, and extensions for ping, transaction, file echoing, bursts, multi streaming, and unidirectional transmission and recording.

5 Referring to Figure 3, a flow chart is depicted in which program code or microprocessor-based microcode in a management function on a user, server, peer, management, or other device 21, 23, or 25 attaches to the network 20 and performs a sequence consisting of probative testing and  
10 analytical steps from which the capacity and utilization of the entire end-to-end system and its component end processor and network parts are stored, displayed, and retrieved , thereby reflecting the past and present condition of the networked system. Furthermore, with user input of proposed  
15 changed conditions, the future capacity, utilization, and performance of the networked system are calculated and reported or displayed.

Referring further to Figure 3, in accordance with the  
20 preferred embodiment of the method of the invention, in step 100 the user signs on through input device 26 to the apparent network speed analysis application (ANSA) 30.



In step 102, the user adjusts default values, if desired. These default values include number of short or long pings and transactions, number of bytes per ping, and time between pings and transactions, as well as whether  
5 ANSA's network calculation is to be based upon a default value for average network message length, a user input value for average network message length, or a value that the system calculates.

10 In step 106, the user enters the IP address (or name) of the target station 24.

In step 108, the user enters the file transfer packet size (MTU).

15 In step 110, the user selects any or all of hop count, duplex, and streaming file transfer cases if throughput analysis is desired.

In step 112, the user selects a desired interactive (transaction response time-oriented) window recommendation routine and/or batch (file transfer) window estimation routine.

Referring to Figure 4, in accordance with a preferred embodiment of the invention, the apparent network speed analysis application (ANSA) 30 performs its throughput routine 34. Responsive to user selection of cases or  
5 functions in step 110, ANSA performs any or all of the functions in steps 132-140.

In step 132, ANSA 30 performs a unidirectional file transfer (FTP) or stream of unidirectional (non-echoed) bursts.

10 In step 134, ANSA 30 performs a bi-directional, concurrent file transfer or stream of echoed bursts.

In step 135, all received bursts are time stamped, and the number of dropped packets determined.

15 In step 136, throughput parameters, to be more thoroughly described hereafter, are calculated.

In step 138, ANSA 30 determines network throughput capacity by calculating such throughput parameters as hop count, duplex, and total throughput factors, as will be described more fully hereafter.

In step 140, network throughput utilization is calculated, as will be more thoroughly described hereafter.

Referring to Figure 5, a geometric chart illustrates various relationships pertinent to the present invention.

5 The Y-axis represents transmission time across the network  
20, so points a, e, and m are successively increasing times.  
The X-axis represents number of bytes per average network  
message, so points a, b, c, and d are successively  
increasing numbers of bytes representing the numbers of  
10 bytes of an average messages under varying conditions in the  
network.

a represents the origin point, which is zero bytes,  
zero time.

The data points (represented by triangles j through q,  
15 and k through r) represent the time it takes probative test  
sample packets to traverse the network. This could be a one  
way transmission or a two way (echo packet or ping)  
transmission.

20 b is the number of bytes in each of the j through q  
test samples.

c is the number of bytes in each of the samples k through r.

n represents the average transmission time for the set of samples j through q, with j representing the best transmission time (no queuing was experienced by this sample), and q represents the time it took for the worst surviving packet to be received.

o represents the average transmission time for the set of samples k through r, with k representing the best transmission time (no queuing was experienced by this sample), and r represents the time it took for the worst surviving packet to be received.

The average time spent waiting on queue in the network before receiving service is represented by segment (m,e). Line segments (e,l) and (m,p) are parallel, because (m,p) remains a constant distance of (m,e) from (e,l).

Chart time (e,a) represents the latency of the network. Certain delays in transmission and receipt across the network occur that are invariant and also independent of the length of the packet being transmitted. The sum total

across the network of such delays is the network's latency. The physical length of the network is always a component of latency, for network length causes propagation delay which is identical for long and short frames. Device latency will  
5 be another component of latency if and only if the device through which packets are passing can handle longer and shorter frames in the same amount of time. Communication link serialization time is dependent on packet length and is, therefore, never a component of latency. And because  
10 latency here is construed to consist of the invariant delays inherent in the network, network queue delay is also never a component of latency.

The dashed line segment (e,i) represents the streaming speed of the network. As described in this patent  
15 application, this speed is calculated by transmitting a burst of (e.g., 10) packets across the network, and calculating time from the receiver's receipt of the first to the last, and dividing the total number of bytes received by the time, and multiplying by 8 to get the bits per second  
20 throughput speed of the network.

Line segment (e,l) represents the discrete speed of the network (also called "datagram speed"). As first described

in Klassen and Silverman, this is computed by considering the number of bytes per small packet  $b$  from number of bytes per large packet  $c$  and multiplying by 8 to get bits. Then the time for transmission and receipt of  $j$  is subtracted  
5 from the time for transmission and receipt of  $k$  to get time. Bits are divided by time to get bits per second discrete speed of the network.

Note, network discrete speed is less than network throughput speed in a multihop network, and equal in a  
10 single hop network. In this chart, speed is inversely proportional to line slope, so discrete speed will always have equal or greater slope than throughput speed.

The hop count value is derived by dividing the throughput speed by the discrete speed. This value is then  
15 multiplied by two if the throughput speed was calculated with echo packets and the network is full duplex. Geometrically, this calculation is represented in the chart by dividing the inverse of the slope of  $(e,i)$  by the inverse of the slope of  $(e,l)$  and multiplying the result by two.

20 The testing method for determining discrete line speed and discrete utilization involves transmission, receipt, and

time stamping of test samples of (at least) two sizes. The geometric chart depicts samples of two sizes, namely smaller sample packets of  $b$  bytes length, and longer packets of  $c$  bytes length.

5           Line segment  $(f,j)$  represents the service time,  $T_s$ , for a packet of  $b$  bytes in this network. Line segment  $(j,n)$  represents the average time waiting for service,  $T_w$ , for a packet of  $b$  bytes in the network. Line segment  $(f,n)$  represents the service time plus the average time, known as  
10   total queue time or  $T_q$ , for a packet of  $b$  bytes in the network.  $T_q = T_s + T_w$ .

          Similarly, line segment  $(g,k)$  represents the service time  $T_s$  for a packet of  $c$  bytes in this network. Line segment  $(k,o)$  represents the time waiting for service,  $T_w$ ,  
15   for a packet of  $c$  bytes in the network. Line segment  $(g,o)$  represents  $T_q$ , the service time plus the average time waiting for service, for a packet of  $c$  bytes.

          The total service time plus the average wait time for a packet is commonly called "total queue time," and is denoted  
20   as  $T_q$ . In the geometric chart, the  $T_q$  value is represented by the perpendicular raised from segment  $(e,h)$  to segment

(m,p) .

Network latency, segment (e,a), is determined from samples j through q and samples k through r as follows. Samples j and k determine the slope of segment (e,l), whose  
5 inverse represents the network's discrete speed. Segment (b,j) represents the total response time for transmission and receipt of a packet of b bytes that experiences no network queueing. Therefore (b bytes \* 8) / network  
discrete speed =  $T_s$  for b bytes, which is segment (f,j) on  
10 the chart. Subtracting  $T_s$  from total response time gives latency; that is,  $(b,j) - (f,j) = (b,f)$ . In this example latency is calculated from discrete speed and short packet response time. Discrete speed and large packet response  
time could similarly be used to calculate latency, as also  
15 could be averaged large and small packet response time and size together with discrete speed.

### ***Probative Packet Burst Transmissions***

In accordance with the preferred embodiment of the invention, in order to establish network characteristics



including the historical, current, and predicted future of  
states of a network for all types of network traffic,  
including interactive, browser, batch, and realtime traffic,  
probative transmissions, including echoed and non-echoed  
5 packets, of like and differing lengths, of like and  
differing network priority, individually and in bursts or  
streams, are sent and transit times measured, and queuing  
theory applied to the results. This involves transmission  
and receipt of bursts of probative packets suitable for  
10 determining the network's streaming speed (its maximum end  
to end speed for file and other batch transfers).

#### ***Burst Test Calculation of Streaming Speed***

In accordance with the preferred embodiment of the  
invention, a method for calculating network streaming speed  
15 based upon bursting is provided. Because the method uses  
bursting, and not streaming, it is neither disruptive to  
other users of the network nor is it affected by other users  
of the network, including large networks and Internet  
connections.

20 Unlike other bursting methods, the method of the

preferred embodiment of the invention does not yield inaccurate results as a result of failure to fill the network. This is because burst timing is performed only at the receiver, the node executing throughput routine 34.

5 This means that the network will slow probative test frames down to the network's bottleneck speed, which is the rate at which they arrive at the receiver. As previously noted, in step 135, the receiver time stamps the receipt of the first through last frames of the burst, checks to make sure that  
10 frames arrived in sequence and without retransmission, and then in steps 136 calculates network speed by dividing the number of bits received in the burst by the total receipt time for the burst. If time stamping of receipt can only be done at the beginning of a frame or at the end of a frame,  
15 but not both, then in a burst of  $n$  frames, the measurements will actually cover  $n-1$  frames, so the measure of bits divided by time will be for  $n-1$  frames. In this manner, for a series of bursts, an accurate measure of best network streaming speed and average network streaming speed is  
20 derived. The best network streaming speed equals the actual clocking rate for the network from an end user application perspective, such as file transfer.

In accordance with this method, in order to use a burst

test to determine the network's streaming speed, a plurality of packets is transmitted and the timestamping of their receipt by the receiver is used to calculate the network's current speed. The best observed speed for the plurality of samples is then taken to be the network's maximum speed. In a small network (such as a test laboratory), conditions may obtain such that traffic competing with the test bursts is not too great in volume, and that testing will stabilize quickly. By this it is meant that during testing it is possible to display the the best speed determined by testing, or for an expert system to keep track of the best speed yet determined by testing. Then subsequent test burst speeds can be observed and compared. The observer (or expert system) can then use various judgments to determine whether testing has stabilized. For example, if no new maximum has been observed after a certain period of time, and if several instances of current streaming speeds close to the maximum have been observed, then that is an indicator that testing has stabilized.

20 In a busy network, or across the internet, achieving testing stabilization may take more testing time than is desirable. This will depend on the traffic characteristics of the network, the number of actual hops in the network,

and the number and size of of packets per burst test. (Note that in the preferred embodiment, a burst consists of ten packets, each packet being of identical size, namely the network's maximum MTU size which is the largest possible packet that can be sent without fragmentation occurring, meaning that this is the largest packet the network can transmit without the packet having to be divided into smaller packets and then reassembled at any point along the network path. It would be possible to transmit a different number of frames per burst, and these frames could be of like or different sizes, some or none of which could be the network's MTU size.) In order to minimize test time while achieving accurate results, the following method can be utilized.

15 In order to achieve accurate streaming speed calculation results with minimum testing, it is possible to create a "logical" best burst receipt constructed from all of the frames in all of the tests. This can be done in any of several ways. Without loss generality, the preferred embodiment in which there are ten uniform packets of maximum MTU length per burst is considered. The logical best burst is constructed by selecting the best receipt time for any packet number one. (Note that the set of bursts under

consideration are only those for which all packets in the burst arrived at the receiver in proper sequence number order with no retransmissions.) Then from all the bursts, the best packets number one and two receipt time is

5 selected, and so on up to packets one through ten. In this manner, a "logical burst", consisting of the best receipt times for  $i$  packets (where  $i = 1, 2, \dots, 10$ ) is constructed starting from a normalized receipt time of zero for the first bit received. The total arrival time of this logical  
10 best burst is calculated and the total number of bits in the burst divided by the receipt time for the burst. This results in an accurate value even in busy, multihop networks (including the internet) even in busy conditions.

An alternative method for calculating the network's  
15 streaming speed using the logical best burst is to compute the average packet receipt time (last bit receipt time minus first bit receipt time averaged for each packet in the logical best burst) and the average interpacket receipt time gap (the "spacing" that the network imposes upon packets  
20 across this connection between individual packets). The bits per packet (maximum MTU) divided by the addition of gap time and packet receipt time will also give the network's streaming speed. Note: if the network is full duplex and if

echo packets are used, then the calculation needs to be performed using the formula:

$$2 * \text{bits per packet} / \text{receipt time}$$

where receipt time = gap time per packet plus first-to-last  
5 bit receipt time per packet.

Note also that an "average burst" can be constructed from the set of test bursts by the same technique as construction of the "best burst" by constructing a burst comprised of the averaged receipt times for all successful  
10 bursts (i.e., no dropped frames, no retransmissions, no out of sequence frames in the burst).

Using the exact same techniques and formulas, the network's average streaming speed can be computed, and then:

$$(1 - (\text{average streaming speed} / \text{streaming speed})) * 100 = \text{percent utilization.}$$
  
15

## Maximum Packet Size Testing

In the preferred embodiment, burst testing is preceded by network maximum packet size testing, a function of routines 38. Packets of successively larger size are sent with "do not fragment" turned on until a network message indicating fragmentation was necessary is received. Then successively smaller packets (or some other of many possible algorithms for selection of length of trial packets) are transmitted until no fragmentation is required and transmission is successful. Once the network's maximum packet size (MTU) is thus determined, bursts of packets of that size are used in steps 130-134 for determining network streaming speed. By this approach the number of interframe gaps and other such overheads are minimized, thus ensuring that the network streaming speed measures the maximum throughput the network and can provide for a perfectly tuned application using the maximum size packets which is accurate across all networks, including the Internet, and non disruptive.

## Average Streaming Utilization

By measuring both best and average network speeds, where the average speed is the speed test traffic was able to achieve in competition with the average amount of user traffic, the method of the preferred embodiment of the invention provides for determining in steps 138 and 140 average streaming utilization of the network. The average streaming utilization percent is determined by taking the ratio of average burst rate to best burst rate and multiplying by 100.

10 In performing burst testing, a certain number of packets will be lost. This is normal in networks and occurs, for example, when some device in the network has full buffers when new frames arrive. In such cases, newly arriving frames are discarded until that device has had the chance to empty sufficient buffer space to accept new frames. Calculation of utilization therefore takes into account such occurrences of frame discards. In accordance with an embodiment of the present invention, frame discards are counted as instances of 100% utilization.

20 As discussed above, in accordance with the preferred embodiment of the present invention, bursting, and not streaming, is used in steps 130-134 as the preferred method of determining the network's streaming speed. This is



because streaming must be performed in an otherwise empty network in order to be accurate and in order not to interfere with others who are using the network for productive purposes.

5

### ***Complex Burst Test Method***

10 In accordance with an alternative embodiment of the present invention, the accuracy of the complex burst test method previously described is improved by ensuring that a sufficient number of short frames are transmitted in a burst so as to fill the network. Five things must be noted here.

15 First, since in the complex burst test method identical numbers of short and long frames are transmitted per test iteration, if the number of frames in the burst of shorter frames will fill the network, then so will the burst of longer frames.

Second, if this test is performed properly and completes properly, it will result in a measure of network speed slightly greater than the speed determined by the

simple burst test. This is because the simple burst test measures the network at the application level, which means that factors that induce small amounts of network dead time, such as interframe gaps, come into play. In the complex  
5 burst test, the network's pure data rate is measured, absolutely exclusive of any dead time. This means that the complex burst test measures the network's clocking rate, while the simple burst test measures the slightly lower rate of maximum possible application-level throughput. As such,  
10 the complex burst test is particularly well suited to measuring the clocking rate of the network from the telco (communications service provider) perspective, while the simple burst test is particularly well suited to measure the network's performance from a streaming user application  
15 perspective.

Third, in a long or fast network, the complex burst test may require a transmission of a large enough number of frames so as to be disruptive to user traffic.

Fourth, the window size, that is, the number of frames  
20 required to fill the network, is determined by using information gathered in other tests (specifically, from best and average round trip discrete time and from simple burst

tests), so if the complex test is run after these two tests have completed, there is no problem in determining the window size required for a proper complex burst run.

5 Fifth, from a practical perspective, the complex burst test is probably best run in large networks, fast networks at times when the network is not busy. Also, the complex burst test is ideally run on isolated, test networks for the purpose of testing vendor equipment capacity, and on newly operational links to verify that they have actually supplied  
10 the bandwidth required.

### **FORMULAS**

In accordance with a specific embodiment of the invention, the explicit formulas calculated by ANSA are set  
15 forth hereafter. By deriving apparent bandwidth (response time capacity), network streaming utilization, network discrete utilization, and the network message size, all of queuing theory is now available for analytic and predictive purposes and can be applied, as described in Klassen and  
20 Silverman.

**Network Streaming Speed** (the rate at which a network is capable of streaming a user's traffic, end to end, when otherwise vacant of traffic)--

Formula:

$$\text{network streaming speed} = \frac{\text{total bits in burst}}{\text{total receipt time for burst}}$$

Notes:

(1) This formula is applied to the best sample burst experienced for which all frames in the burst successfully arrived in sequence and with no retransmissions.

(2) Total receipt time for burst = receipt time for last bit in burst - receipt time for first bit in burst.

(3) In the preferred embodiment, burst testing that determines the network's streaming speed is preceded by network maximum packet size testing. Once the network's maximum MTU (maximum packet size) is determined, bursts of packets of that size are used in testing to determine

network's streaming speed. This ensures that the network streaming speed measures the maximum throughput the network can provide for a perfectly tuned application using the maximum size packets. By this method the number of

5 interframe gaps and other such overheads are absolutely minimized.

(4) The network streaming speed is the true measure of the slowest hop in the actual network.

**Average Network Streaming Speed** (the speed at which the

10 network is capable of streaming a user's traffic under current conditions, including the traffic of other users of the network.)--

Formula:

average network streaming speed = total bits in burst /

15 average total receipt time for burst

Notes:

(1) In testing, a plurality of bursts of packets of a given

size are transmitted. To calculate the network's average streaming speed, the average receipt time is calculated over the entire set of burst samples in the test, and this average time is what is used in the denominator of the formula.

(2) During the course of testing, packets will be lost and will require retransmission. These retransmissions will occur and will increase the total receipt time. Since the retransmissions are accounted for in the total receipt time, retransmissions are implicitly accounted for in the calculation of average network streaming speed.

(3) Below, this value is also referred to as "current network streaming speed".

**Network Streaming Utilization** (the percent of the network currently in use that is not available to a user commencing use of an application that will stream traffic along the end to end network connection.)--

Formula:

network streaming utilization = (network streaming  
speed / average network streaming speed) \* 100.

As noted above, this includes overhead incurred for  
retransmissions.

5

**Network Discrete Speed** (this is the end to end rate at which  
datagram traffic will be carried across the network if the  
network is empty of other traffic and there is no  
retransmission required.)

10 Formula:

network discrete speed = (long packet bits - short  
packet bits) / (best long time - best short time)

15 **Network Queue Wait Time, "Tw,"** as derived from testing (Tw  
is the actual average wait time experienced on queues before  
obtaining service between the ANSA test station and the  
target station)--

Formula:

$$Tw = S (\text{ping sample } i \text{ time} - \text{best ping sample time})$$
  
for all  $i$  / total number of ping samples.

That is, divide the sum of each ping sample minus the best sample by the total number of samples. This gives the  
5 average cross-network wait time for service.

**Standard Deviation Network Queue Wait Time,  $\sigma Tw$**  (the mean statistical variation of wait times on network queues prior to receiving service )--

Formula:

10 The standard square of the square root of the differences between sample times, well known in statistical analysis.

Notes

(1) Standard statistical analysis is applied to ANSA  
15 probative test sample packets to determine the standard deviation and average wait times for service. Subtraction of the time of best transmission from all worse sample



transmission times is used to derive queue wait time values. This works because the packet samples in a single set of transmissions are of equal length. This means that subtraction of best round trip or one way time from any other round trip or one way time eliminates device latency time, propagation delay time, and service time, which are equal for all samples. The subtraction therefore identifies all and only the time spent waiting for service, and thus  $T_w$  and  $\sigma T_w$  are derived from test packet samples by use of the formulas above.

(2) The formulas above for deriving values for  $T_w$  and  $\sigma T_w$  are applied to ANSA test packets. As such, they represent empirical measures of the actual current state of the network. The formulas and ratios described below provide a queuing theory-based analytic means for using these empirical measurements as a basis for determining the utilization and message size characteristics of the network.

**Network Queue Wait Time, " $T_w$ ," from a queueing theory perspective** (the queueing theory formula for  $T_w$ ):

Formula:

$$T_w = pT_s / (1-p)$$

where  $T_s$  is the service time for an average network packet and  $p$  is the network's utilization.

**Standard Deviation Network Queue Wait Time,  $\sigma T_w$ , from a**

5 **queueing theory perspective** (the queueing theory formula for  $\sigma T_w$  ):

Formula:

$$\sigma T_w = [\sqrt{p * (2 - p) } ] * (T_s / (1 - p))$$

In accordance with this embodiment of the invention, this  
 10 formula is used in derivation of the value of a network's  
 discrete utilization value. The specific formulas in which  
 it is used are shown in the copending application, S/N  
 \_\_\_\_\_. The importance of this queueing theory formula is  
 that it relates the empirically derived value for  $\sigma T_w$  to  
 15 utilization and service time values and, therefore, provides  
 a necessary link between ANSAT testing methods and the  
 derivation of discrete utilization and average network  
 message length.

### Advantages over the Prior Art

It is an advantage of the invention that there is provided an improved system and method for network measurement and planning.

5        It is an advantage of the invention that there is provided an improved method for calculating the streaming speed of a network.

10       It is an advantage of the invention that there is provided an improved method for determining end to end network streaming utilization.

      It is an advantage of the invention that there is provided an improved method for tuning a network.

15       It is an advantage of the invention that there is provided an improved method for estimating the current performance of a network and its future performance under changed conditions ("what-if" scenarios).

## Alternative Embodiments

It will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without departing from the spirit and scope of the invention. In particular, it is within the scope of the invention to provide a computer program product or program element, or a program storage or memory device such as a solid or fluid transmission medium, magnetic or optical wire, tape or disc, or the like, for storing signals readable by a machine, for controlling the operation of a computer according to the method of the invention and/or to structure its components in accordance with the system of the invention.

Further, each step of the method may be executed on any general computer, such as an IBM System 390, AS/400, PC or the like and pursuant to one or more, or a part of one or more, program elements, modules or objects generated from any programming language, such as C++, Java, Pl/1, or the like. And still further, each said step, or a file or object or the like implementing each said step, may be

executed by special purpose hardware or a circuit module  
designed for that purpose.

Accordingly, the scope of protection of this invention  
5 is limited only by the following claims and their  
equivalents.

END9 2000 0102 US1